

Serial Number: 09/645,705

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

In the Claims

Please amend the claims as follows:

1. (Currently Amended) A first synthetic nucleic acid molecule comprising at least 300 nucleotides of a coding region for a reporter polypeptide which has at least 90% amino acid sequence identity to a reporter polypeptide encoded by a wild type nucleic acid sequence, wherein the codon composition of the first synthetic nucleic acid molecule is different at more than 25% of the codons from that of the wild type nucleic acid sequence and is different than the codon composition of a second synthetic nucleic acid molecule which encodes a reporter polypeptide which has at least 90% amino acid sequence identity to the reporter polypeptide encoded by the wild type nucleic acid sequence, wherein the codons in the second synthetic nucleic acid molecule that are different than the codons in the wild type nucleic acid sequence are mammalian high usage codons selected to result in the second synthetic nucleic acid molecule having a reduced number of a combination of different mammalian transcription factor binding sequences, and optionally a reduced number of intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences relative to the wild type nucleic acid sequence, wherein the codons which differ in the first synthetic nucleic acid molecule relative to the second synthetic nucleic acid molecule are mammalian codons selected to result in the first synthetic nucleic acid molecule having a reduced number of a combination of different mammalian transcription factor binding sequences, and optionally a reduced number of intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences, that are introduced to the second synthetic nucleic acid molecule by selecting the mammalian high usage codons, wherein the mammalian transcription factor binding sequences are those present in a database of transcription factor binding sequences, wherein the wild type nucleic acid sequence and the first synthetic nucleic acid molecule encode ~~encodes~~ chloramphenicol acetyltransferase, *Renilla* luciferase, beetle luciferase, beta-lactamase, beta-glucuronidase or beta-galactosidase, wherein the first synthetic nucleic acid molecule has at least 3-fold fewer of the combination of mammalian transcription factor binding sequences and optionally intron splice sites, poly(A)

Serial Number: 09/645,705

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

addition sites or prokaryotic 5' noncoding regulatory sequences relative to those in the wild-type nucleic acid sequence.

2. (Canceled).
3. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the codon composition of the first synthetic nucleic acid molecule differs from the wild type nucleic acid sequence at more than 35% of the codons.
4. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the codon composition of the first synthetic nucleic acid molecule differs from the wild type nucleic acid sequence at more than 45% of the codons.
5. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the codon composition of the first synthetic nucleic acid molecule differs from the wild type nucleic acid sequence at more than 55% of the codons.
- 6-8. (Canceled).
9. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 which encodes a luciferase.
10. (Canceled).
11. (Previously Presented) The first synthetic nucleic acid molecule of claim 9 wherein the wild type nucleic acid sequence encodes a beetle luciferase.
12. (Previously Presented) The first synthetic nucleic acid molecule of claim 11 wherein the first synthetic nucleic acid molecule encodes the amino acid valine at position 224.

Serial Number: 09/645,705

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

13-14. (Canceled).

15. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 or 9 wherein the majority of codons which differ in the second synthetic nucleic acid molecule are those which are preferred codons in humans.

16-17. (Canceled).

18. (Currently Amended) A synthetic nucleic acid molecule comprising SEQ ID NO:7 (GRver5), SEQ ID NO:8 (GRver6), SEQ ID NO:9 (GRver5.1), or SEQ ID NO:297 (GRver5.1), or a nucleic acid molecule which is capable of hybridizing thereto under high stringency conditions, or the complement of the hybridizable nucleic acid molecule which encodes a luciferase, wherein the complement of the hybridizable nucleic acid molecule has at least 3-fold fewer of a combination of mammalian transcription factor binding sequences and optionally intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences relative to a wild-type nucleotide sequence encoding the luciferase.

19. (Canceled).

20. (Previously Presented) The first synthetic nucleic acid molecule of claim 15 wherein the majority of codons which differ are the human codons CGC, CTG, TCT, AGC, ACC, CCA, CCT, GCC, GGC, GTG, ATC, ATT, AAG, AAC, CAG, CAC, GAG, GAC, TAC, TGC and TTC.

21. (Previously Presented) The first synthetic nucleic acid molecule of claim 15 wherein the majority of codons which differ are the human codons CGC, CTG, TCT, ACC, CCA, GCC, GGC, GTC, and ATC or codons CGT, TTG, AGC, ACT, CCT, GCT, GGT, GTG and ATT.

Serial Number: 09/645,705

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

22-23. (Canceled).

24. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the first synthetic nucleic acid molecule is expressed in a mammalian host cell at a level which is greater than that of the wild type nucleic acid sequence.

25. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the first synthetic nucleic acid molecule has an increased number of CTG or TTG leucine-encoding codons.

26. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the first synthetic nucleic acid molecule has an increased number of GTG or GTC valine-encoding codons.

27. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the first synthetic nucleic acid molecule has an increased number of GGC or GGT glycine-encoding codons.

28. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the first synthetic nucleic acid molecule has an increased number of ATC or ATT isoleucine-encoding codons.

29. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the first synthetic nucleic acid molecule has an increased number of CCA or CCT proline-encoding codons.

30. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the first synthetic nucleic acid molecule has an increased number of CGC or CGT arginine-encoding codons.

Serial Number: 09/645,705

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

31. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the first synthetic nucleic acid molecule has an increased number of AGC or TCT serine-encoding codons.
32. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the first synthetic nucleic acid molecule has an increased number of ACC or ACT threonine-encoding codons.
33. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the first synthetic nucleic acid molecule has an increased number of GCC or GCT alanine-encoding codons.
34. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the codons in the first synthetic nucleic acid molecule which differ encode the same amino acids as the corresponding codons in the wild type nucleic acid sequence.
35. (Previously Presented) A plasmid comprising the first synthetic nucleic acid molecule of claim 1.
36. (Previously Presented) An expression vector comprising the first synthetic nucleic acid molecule of claim 1 linked to a promoter functional in a cell.
37. (Previously Presented) The expression vector of claim 36 wherein the first synthetic nucleic acid molecule is operatively linked to a Kozak consensus sequence.
38. (Original) The expression vector of claim 36 wherein the promoter is functional in a mammalian cell.

Serial Number: 09/645,705

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

39. (Original) The expression vector of claim 36 wherein the promoter is functional in a human cell.
40. (Canceled).
41. (Original) The expression vector of claim 36 wherein the expression vector further comprises a multiple cloning site.
42. (Previously Presented) The expression vector of claim 41 wherein the expression vector comprises a multiple cloning site positioned between the promoter and the first synthetic nucleic acid molecule.
43. (Previously Presented) The expression vector of claim 41 wherein the expression vector comprises a multiple cloning site positioned downstream from the first synthetic nucleic acid molecule.
44. (Previously Presented) An isolated host cell comprising the expression vector of claim 36.
45. (Previously Presented) A kit comprising, in suitable container means, the expression vector of claim 36, wherein the first synthetic nucleic acid molecule encodes a reporter molecule.
46. (Canceled).
47. (Currently Amended) A first polynucleotide which hybridizes under high ~~medium~~ stringency hybridization conditions to SEQ ID NO:9 (GRver5.1), SEQ ID NO:18 (RD156-1H9), SEQ ID NO:297 (GRver5.1), SEQ ID NO:301 (RD156-1H9), or the complement thereof, and comprises an open reading frame encoding a beetle luciferase polypeptide which has at least 90% amino acid sequence identity to a luciferase having SEQ ID NO:23 encoded by a

Serial Number: 09/645,705

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

corresponding wild type nucleic acid sequence having SEQ ID NO:1, wherein the codon composition of the open reading frame of the first polynucleotide is different at more than 25% of the codons from that of the wild type luciferase nucleic acid sequence and is different than the codon composition of a second polynucleotide which encodes a polypeptide which has at least 90% amino acid sequence identity to the polypeptide encoded by the wild type nucleic acid sequence, wherein the codons in the second polynucleotide that are different than the codons in the wild type nucleic acid sequence are mammalian high usage codons selected to result in the second polynucleotide having a reduced number of a combination of different mammalian transcription factor binding sequences, intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences relative to the wild type nucleic acid sequence, wherein the codons which differ in the first polynucleotide relative to the second polynucleotide are mammalian codons selected to result in the open reading frame in the first polynucleotide having a reduced number of a combination of different mammalian transcription factor binding sequences, and optionally a reduced number of intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences, that are introduced to the second polynucleotide by selecting the mammalian high usage codons, wherein the mammalian transcription factor binding sequences are those present in a database of transcription factor binding sequences, wherein the first polynucleotide has at least 3-fold fewer of the combination of different mammalian transcription factor binding sequences, intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences relative to the wild type nucleic acid sequence.

48-59. (Canceled).

60. (Previously Presented) The first synthetic nucleic acid molecule of claim 1 wherein the first synthetic nucleic acid molecule is expressed at a level which is at least 110% of that of the wild type nucleic acid sequence in a cell or cell extract under identical conditions.

61-63. (Canceled).

Serial Number: 09/645,705

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

64. (Withdrawn) The vector of claim 63 wherein the synthetic nucleic acid molecule does not encode a polypeptide.

65-66. (Canceled).

67. (Currently Amended) A first synthetic nucleic acid molecule comprising at least 300 nucleotides of a coding region for a luciferase which has at least 90% amino acid sequence identity to a reporter polypeptide encoded by a wild type beetle luciferase nucleic acid sequence, wherein the codon composition of the first synthetic nucleic acid molecule is different at more than 25% of the codons from that of the wild type nucleic acid sequence and is different than the codon composition of a second synthetic nucleic acid molecule which encodes a luciferase which has at least 90% amino acid sequence identity to the luciferase encoded by the wild type nucleic acid sequence, wherein the codons in the second synthetic nucleic acid molecule that are different than the codons in the wild type nucleic acid sequence are mammalian high usage codons selected to result in the second synthetic nucleic acid molecule having a reduced number of a combination of different mammalian transcription factor binding sequences, and optionally a reduced number of intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences relative to the wild type nucleic acid sequence, wherein the codons which differ in the first synthetic nucleic acid molecule relative to the second synthetic nucleic acid molecule are mammalian codons selected so as to result in the first synthetic nucleic acid molecule having a reduced number of a combination of different mammalian transcription factor binding sequences, and optionally a reduced number of intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences, that are introduced to the second synthetic nucleic acid molecule by selecting the mammalian high usage codons, wherein the mammalian transcription factor binding sequences are those present in a database of transcription factor binding sequences, wherein the first synthetic polynucleotide has at least 3-fold fewer of the combination of different mammalian transcription factor binding sequences, intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences relative to the wild type nucleic acid sequence.

Serial Number: 09/645,705

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

68. (Canceled).

69. (Previously Presented) The first synthetic nucleic acid molecule of claim 11 or 67 which has 74% or less nucleic acid sequence identity to the wild type nucleic acid sequence.

70. (Previously Presented) The first synthetic nucleic acid molecule of claim 11 or 67 which has at least 40-fold increased expression relative to the wild type nucleic acid sequence.

71. (Previously Presented) The first polynucleotide of claim 47 which hybridizes under high stringency hybridization conditions to SEQ ID NO:9 (GRver5.1), SEQ ID NO:18 (RD156-1H9), SEQ ID NO:297(GRver5.1), SEQ ID NO:301 (RD156-1H9), or the complement thereof.

72-73. (Canceled).

74. (Currently Amended) A first synthetic nucleic acid molecule comprising at least 300 nucleotides of a coding region for a luciferase which has at least 90% amino acid sequence identity to a luciferase encoded by a parent nucleic acid sequence having SEQ ID NO:2, wherein the codon composition of the synthetic nucleic acid molecule is different at more than 25% of the codons from that of the parent nucleic acid sequence and is different than the codon composition of a second synthetic nucleic acid molecule which encodes a luciferase which has at least 90% amino acid sequence identity to the luciferase encoded by the parent nucleic acid sequence, wherein the codons in the second synthetic nucleic acid molecule that are different than the codons in the parent nucleic acid sequence are mammalian high usage codons selected to result in the second synthetic nucleic acid molecule having a reduced number of a combination of different mammalian transcription factor binding sequences, and optionally a reduced number of intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences relative to the parent nucleic acid sequence, wherein the codons which differ in the first synthetic nucleic acid molecule relative to the second synthetic nucleic acid molecule are mammalian

Serial Number: 09/645,705

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

codons selected to result in the first synthetic nucleic acid molecule having a reduced number of a combination of different mammalian transcription factor binding sequences, and optionally a reduced number of intron splice sites, poly(A) addition sites or or prokaryotic 5' noncoding regulatory sequences, that are introduced to the second synthetic nucleic acid molecule by selecting the mammalian high usage codons, wherein the mammalian transcription factor binding sequences are those present in a database of transcription factor binding sequences, wherein the first polynucleotide has at least 3-fold fewer of the combination of different mammalian transcription factor binding sequences, intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences relative to the parent nucleic acid sequence.

75. (Canceled).

76. (Previously Presented) The first synthetic nucleic acid molecule of claim 74 wherein the polypeptide encoded by the first synthetic nucleic acid molecule has at least 95% amino acid identity to the luciferase encoded by the parent nucleic acid sequence.

77. (Previously Presented) The first synthetic nucleic acid molecule of claim 74 which has 74% or less nucleic acid sequence identity to the parent nucleic acid sequence.

78. (Currently Amended) A first polynucleotide which hybridizes under high ~~medium~~ stringency hybridization conditions to SEQ ID NO:9 (GRver5.1) or SEQ ID NO:297 (GRver5.1), or the complement thereof, and comprises an open reading frame encoding a luciferase polypeptide which has at least 90% amino acid sequence identity to a luciferase encoded by a parent nucleic acid sequence having SEQ ID NO:2, wherein the codon composition of the open reading frame of the first polynucleotide is different at more than 25% of the codons from that of the parent nucleic acid sequence and is different than the codon composition of a second polynucleotide which encodes a polypeptide which has at least 90% amino acid sequence identity to the luciferase encoded by the parent nucleic acid sequence, wherein the codons in the second polynucleotide that are different than the codons in the parent nucleic acid sequence are

Serial Number: 09/645,705

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

mammalian high usage codons selected to result in the second polynucleotide having a reduced number of a combination of different mammalian transcription factor binding sequences, and optionally a reduced number of intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences relative to the parent nucleic acid sequence, wherein the codons which differ in the first polynucleotide relative to the second polynucleotide are mammalian codons selected to result in the first polynucleotide having a reduced number of a combination of different mammalian transcription factor binding sequences, and optionally a reduced number of intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences that are introduced to the second polynucleotide by selecting the mammalian high usage codons, wherein the mammalian transcription factor binding sequences are those present in a database of transcription factor binding sequences, wherein the first polynucleotide has at least 3-fold fewer of the combination of different mammalian transcription factor binding sequences, intron splice sites, poly(A) addition sites or prokaryotic 5' noncoding regulatory sequences relative to the parent nucleic acid sequence.

79. (Canceled).

80. (Previously Presented) The first polynucleotide of claim 78 wherein the polypeptide encoded by the first polynucleotide has at least 95% amino acid identity to the luciferase encoded by the parent nucleic acid molecule.

81. (Previously Presented) The first synthetic nucleic acid molecule of claim 1, 67 or 74 wherein the transcription factor binding sequence is at least 5 bases in length.

82. (Previously Presented) The first polynucleotide of claim 47 or 78 wherein the transcription factor binding sequence is at least 5 bases in length.

83-85. (Canceled).

Serial Number: 09/645,705
Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

86. (Previously Presented) The first synthetic sequence of claim 1 wherein the selection of mammalian high usage codons and mammalian codons also reduces the number of restriction endonuclease sites.

87. (Previously Presented) The first polynucleotide of claim 47 or 78 wherein the selection of mammalian high usage codons and mammalian codons also reduces the number of restriction endonuclease sites.

88. (Previously Presented) The first synthetic nucleic acid molecule of claim 67 or 74 wherein the selection of mammalian high usage codons and mammalian codons also reduces the number of restriction endonuclease sites.

89. (Canceled).

90. (Currently Amended) The first synthetic nucleic acid molecule of claim 1 wherein the ~~mammalian transcription factor binding sequences, intron splice sites, poly(A) addition sites and prokaryotic 5' noncoding regulatory sequences in the wild type nucleic acid sequence or the second synthetic nucleic acid sequence are identified with software, wherein the identified intron splice sites are selected from site is AGGTRAGT, AGGTRAG, GGTRAGT or YNCAGG, the identified poly(A) addition sites have site is AATAAA, or the identified prokaryotic 5' noncoding regulatory sequences are selected from sequence is TATAAT, or AGGA or GGAG if a methionine codon is within 12 bases 3' of the AGGA or GGAG, and the identified mammalian transcription factor binding sequences are in a database of transcription factor binding sequences, mutant transcription factor binding sequences and consensus transcription factor binding sequences, and identified under parameters that allow for partial ambiguity with sequences in the database, wherein the codons are selected to reduce the number of identified sequences or sites, and wherein the first synthetic nucleic acid molecule has fewer mammalian transcription factor binding sequences than the second synthetic nucleic acid molecule which has fewer mammalian transcription factor binding sequences than the wild type nucleic acid sequence.~~

Serial Number: 09/645,705

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

91. (Canceled).

92. (Previously Presented) A first synthetic nucleic acid molecule comprising at least 300 nucleotides of a coding region for a reporter polypeptide which has at least 90% amino acid sequence identity to a reporter polypeptide encoded by a wild type nucleic acid sequence, wherein the first synthetic nucleic acid molecule is prepared by replacing codons in the wild type nucleic acid molecule with mammalian high usage codons, yielding a second synthetic nucleic acid molecule, and replacing codons in the second synthetic nucleic acid molecule with mammalian codons selected to reduce the number of a combination of different, known mammalian transcription factor binding sites, yielding the first synthetic nucleic acid molecule, wherein the codon composition of the first synthetic nucleic acid molecule is different at more than 25% of the codons from that of the wild type nucleic acid sequence, wherein the wild type nucleic acid sequence encodes chloramphenicol acetyltransferase, *Renilla* luciferase, beetle luciferase, beta-lactamase, beta-glucuronidase or beta-galactosidase.

93. (Currently Amended) The first synthetic nucleic acid molecule of claim [[91 or]] 92 which has at least 2-fold fewer mammalian transcription factor binding sequences relative to the wild type nucleic acid sequence.

94. (Currently Amended) The first synthetic nucleic acid molecule of claim [[91 or]] 92 wherein codon selection also reduces the number of intron splice sites, poly(A) addition sites or promoter sequences, or a combination thereof.

95. (Currently Amended) The first synthetic nucleic acid molecule of claim 67 wherein the ~~mammalian transcription factor binding sequences, intron splice sites, poly(A) addition sites and prokaryotic 5' noncoding regulatory sequences in the wild type nucleic acid sequence or the second synthetic nucleic acid sequence are identified with software, wherein the identified intron splice sites are selected from~~ site is AGGTRAGT, AGGTRAG, GGTRAGT or YNCAGG, the

Serial Number: 09/645,705

Filing Date: August 24, 2000

Title: SYNTHETIC NUCLEIC ACID MOLECULE COMPOSITIONS AND METHODS OF PREPARATION

~~identified poly(A) addition sites have site is AATAAA, or the identified prokaryotic 5' noncoding regulatory sequences are selected from sequence is TATAAT, or AGGA or GGAG if a methionine codon is within 12 bases 3' of the AGGA or GGAG, and the identified mammalian transcription factor binding sequences are in a database of transcription factor binding sequences, mutant transcription factor binding sequences and consensus transcription factor binding sequences, and identified under parameters that allow for partial ambiguity with sequences in the database, wherein the codons are selected to reduce the number of identified sequences or sites, and wherein the first synthetic nucleic acid molecule has fewer mammalian transcription factor binding sequences than the second synthetic nucleic acid molecule which has fewer mammalian transcription factor binding sequences than the wild type nucleic acid sequence.~~

96. (Currently Amended) The first synthetic nucleic acid molecule of claim 74 wherein the ~~mammalian transcription factor binding sequences, intron splice sites, poly(A) addition sites and prokaryotic 5' noncoding regulatory sequences in the parent nucleic acid sequence or the second synthetic nucleic acid sequence are identified with software, wherein the identified intron splice sites are selected from site is AGGTRAGT, AGGTRAG, GGTRAGT or YNCAGG, the identified poly(A) addition sites have site is AATAAA, or the identified prokaryotic 5' noncoding regulatory sequences are selected from sequence is TATAAT, or AGGA or GGAG if a methionine codon is within 12 bases 3' of the AGGA or GGAG, and the identified mammalian transcription factor binding sequences are in a database of transcription factor binding sequences, mutant transcription factor binding sequences and consensus transcription factor binding sequences, and identified under parameters that allow for partial ambiguity with sequences in the database, wherein the codons are selected to reduce the number of identified sequences or sites, and wherein the first synthetic nucleic acid molecule has fewer mammalian transcription factor binding sequences than the second synthetic nucleic acid molecule which has fewer mammalian transcription factor binding sequences than the parent nucleic acid sequence.~~